

UNIFORM ASYMPTOTICS FOR NONPARAMETRIC QUANTILE REGRESSION WITH AN APPLICATION TO TESTING MONOTONICITY

SOKBAE LEE^{1,2}, KYUNGCHUL SONG³, AND YOON-JAE WHANG¹

ABSTRACT. In this paper, we establish a uniform error rate of a Bahadur representation for local polynomial estimators of quantile regression functions. The error rate is uniform over a range of quantiles, a range of evaluation points in the regressors, and over a wide class of probabilities for observed random variables. Most of the existing results on Bahadur representations for local polynomial quantile regression estimators apply to the fixed data generating process. In the context of testing monotonicity where the null hypothesis is of a complex composite hypothesis, it is particularly relevant to establish Bahadur expansions that hold uniformly over a large class of data generating processes. In addition, we establish the same error rate for bootstrap local polynomial estimators which can be useful for various bootstrap inference. As an illustration, we apply to testing monotonicity of quantile regression and present Monte Carlo experiments based on this example.

KEY WORDS. Bootstrap, conditional moment inequalities, kernel estimation, local polynomial estimation, L_p norm, nonparametric testing, Poissonization, quantile regression, testing monotonicity, uniform asymptotics

¹DEPARTMENT OF ECONOMICS, SEOUL NATIONAL UNIVERSITY, 1 GWANAK-RO, GWANAK-GU, SEOUL, 151-742, REPUBLIC OF KOREA.

²CENTRE FOR MICRODATA METHODS AND PRACTICE, INSTITUTE FOR FISCAL STUDIES, 7 RIDGMOUNT STREET, LONDON, WC1E 7AE, UK.

³VANCOUVER SCHOOL OF ECONOMICS, UNIVERSITY OF BRITISH COLUMBIA, 997 - 1873 EAST MALL, VANCOUVER, BC, V6T 1Z1, CANADA

E-mail addresses: sokbae@snu.ac.kr, kysong@mail.ubc.ca, whang@snu.ac.kr.

Date: August 27, 2015.

Lee's work was supported by the European Research Council (ERC-2009-StG-240910-ROMETA) and by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2012S1A5A8023573). Song acknowledges the financial support of Social Sciences and Humanities Research Council of Canada. Whang's work was supported by the SNU Creative Leading Researcher Grant.

1. INTRODUCTION

In this paper, we establish a Bahadur representation of a local polynomial estimator of a nonparametric quantile regression function that is uniform over a range of quantiles, a range of evaluation points in the regressors, and a wide class of probabilities underlying the distributions of observed random variables. We also establish a Bahadur representation for a bootstrap estimator of a nonparametric quantile regression function.

There are several existing results of Bahadur representation of the local polynomial quantile regression estimator in the literature. Chaudhuri (1991) is the classical result on a local polynomial quantile regression estimator with a uniform kernel. His result is pointwise in that the representation holds at one quantile, for a fixed point, and for a given data generating process. For recent contributions that are closely related to this paper, see Kong, Linton, and Xia (2010), Guerre and Sabbah (2012), Kong, Linton, and Xia (2013), and Qu and Yoon (2015), among others. Kong, Linton, and Xia (2010) obtain a Bahadur representation for a local polynomial M-estimator, including the quantile regression estimator as a special case, for strongly mixing stationary processes. Their result holds uniformly for a range of evaluation points in the regressors but at a fixed quantile for a given data generating process. Guerre and Sabbah (2012) obtain Bahadur representations that hold uniformly over a range of quantiles, a range of evaluation points in the regressors, and a range of bandwidths for independent and identically distributed (i.i.d.) data. However, their result is for a fixed data generating process. Kong, Linton, and Xia (2013) extend to the case when the dependent variable is randomly censored for i.i.d. data and obtain the representation that is uniform over the evaluation points. Qu and Yoon (2015) consider estimating the conditional quantile process nonparametrically for the i.i.d. data using local linear regression, with a focus on quantile monotonicity. They develop a Bahadur representation that is uniform in the quantiles but at a fixed evaluation point for a given data generating process. It seems that our work is the first that obtains a Bahadur representation that holds uniformly over the triple: the quantile, the evaluation point, and the underlying distribution. However, our result is for a fixed bandwidth, unlike Guerre and Sabbah (2012).

The most distinctive feature of our result is that the Bahadur representation is uniform over a wide class of probabilities. Uniformity of asymptotic approximation in probabilities has long drawn interest in statistical decision theory and empirical process theory. Uniformity in asymptotic approximation is generally crucial for procuring finite sample stability of size or coverage probability in inference. See Andrews, Cheng, and Guggenberger (2011) for its emphasis and general tools for uniform asymptotic results. In the recent literature of econometrics, identifying the class of probabilities for which the uniformity holds, and their plausibility in practice, have received growing attention, along with increasing interests in

models based on inequality restrictions. See, for example, Andrews and Soares (2010) and Andrews and Shi (2013) among many others.

To see the issue of uniformity, consider a simple testing problem:

$$(1.1) \quad H_0 : \forall x \in \mathcal{X}, \frac{\partial \text{Med}[Y|X=x]}{\partial x} \geq 0 \text{ vs. } H_1 : \exists x \in \mathcal{X}, \frac{\partial \text{Med}[Y|X=x]}{\partial x} < 0,$$

where $\text{Med}[Y|X=x]$ is the conditional median of Y given $X=x$ and \mathcal{X} is a region of interest. Then one may develop a nonparametric test statistic using the local polynomial quantile regression estimator (e.g. the L_p statistic as in Section 3). The behavior of this nonparametric test statistic depends crucially on the contact set $B := \{x \in \mathcal{X} : \partial \text{Med}[Y|X=x]/\partial x = 0\}$. To emphasize the issue of uniformity, consider a sequence of data generating processes indexed by n . For example, we take the sequence of the true conditional median functions to be $\text{Med}_n[Y|X=x] = x^3/n$ on $\mathcal{X} = [-1, 1]$. Then for each n , the corresponding contact set is a singleton set, that is $B_n = \{0\}$; however, $\text{Med}_n[Y|X=x] = x^3/n$ converges to 0 uniformly in $x \in \mathcal{X}$ as $n \rightarrow \infty$. In other words, as n gets large, the true function looks flat on \mathcal{X} , but the population contact set is always the singleton set at zero for each n . This suggests that the pointwise asymptotic theory may not be adequate for finite sample approximation. Therefore, it is important to develop uniform asymptotics for the local polynomial quantile regression estimator by establishing the Bahadur representation that is uniform over a large class of probabilities.

We illustrate the usefulness of our Bahadur representation by applying it to testing monotonicity of quantile regression that includes (1.1) as a special case. Our proposed test uses the framework of Lee, Song, and Whang (2015, LSW hereafter). They provide a general method of testing inequality restrictions for nonparametric functions, and make use of this paper's result in establishing sufficient conditions for one of their results.

The remainder of the paper is as follows. Section 2 presents the main results of the paper, Section 3 gives an application of our main results in the context of testing monotonicity, Section 4 concludes, and Section 5 gives all the proofs.

2. UNIFORM ASYMPTOTICS

This section provides uniform Bahadur representations for local polynomial quantile regression estimators and considers their bootstrap version as well.

2.1. Uniform Bahadur Representation for Local Polynomial Quantile Regression Estimators. In this subsection, we present a Bahadur representation of a local polynomial quantile regression estimator that can be useful for a variety of purposes.

Let $(B^\top, X^\top, L)^\top$, with $B \equiv (B_1, \dots, B_{\bar{L}})^\top \in \mathbf{R}^{\bar{L}}$, and $X \in \mathbf{R}^d$, be a random vector such that the joint distribution of $(B^\top, X^\top)^\top$ is absolutely continuous with respect to Lebesgue

measure and L is a discrete random variable taking values from $\mathbb{N}_L \equiv \{1, 2, \dots, \bar{L}\}$. For each $x \in \mathbf{R}^d$ and $k \in \mathbb{N}_L$, we assume that the conditional distribution of B_l given $(X, L) = (x, k)$ is the same across $l = 1, \dots, k$. It is unconventional to consider a vector B , but it is useful to do so here to cover the case where we observe multiple outcomes from the same conditional distribution.

Let $q_k(\tau|x)$ denote the τ -th quantile of B_l conditional on $X = x$ and $L = k$, where $\tau \in (0, 1)$. That is, $P\{B_l \leq q_k(\tau|x) | X = x, L = k\} = \tau$ for all x in the support of X and all $k \in \{1, \dots, \bar{L}\}$. We write

$$B_l = q_k(\tau|X) + \varepsilon_{\tau,lk}, \quad \tau \in (0, 1), \quad \text{for all } k \in \{1, \dots, \bar{L}\},$$

where $\varepsilon_{\tau,lk}$ is a continuous random variable such that the τ -th conditional quantile of $\varepsilon_{\tau,lk}$ given X and $L = k$ is equal to zero.

Suppose that we are given a random sample $\{(B_i^\top, X_i^\top, L_i)^\top\}_{i=1}^n$ of $(B^\top, X^\top, L)^\top$.¹ Assume that $q_k(\tau|x)$ is $(r+1)$ -times continuously differentiable with respect to x , where $r \geq 1$. We use a local polynomial estimator $\hat{q}_k(\tau|x)$, similar to Chaudhuri (1991). For $u \equiv (u_1, \dots, u_d)$, a d -dimensional vector of nonnegative integers, let $[u] = u_1 + \dots + u_d$. Let A_r be the set of all d -dimensional vectors u such that $[u] \leq r$, and let $|A_r|$ denote the number of elements in A_r . For $z = (z_1, \dots, z_d)^\top \in \mathbf{R}^d$ with $u = (u_1, \dots, u_d)^\top \in A_r$, let $z^u = \prod_{m=1}^d z_m^{u_m}$. Now define $c(z) = (z^u)_{u \in A_r}$, for $z \in \mathbf{R}^d$. Note that $c(z)$ is a vector of dimension $|A_r|$. For $u = (u_1, \dots, u_d)^\top \in A_r$, and $r+1$ times differentiable map f on \mathbf{R}^d , we define the following derivative:

$$(D^u f)(x) \equiv \frac{\partial^{[u]}}{\partial x_1^{u_1} \cdots \partial x_d^{u_d}} f(x),$$

where $[u] = u_1 + \dots + u_d$. Then we define $\gamma_{\tau,k}(x) \equiv (\gamma_{\tau,k,u}(x))_{u \in A_r}$, where

$$\gamma_{\tau,k,u}(x) \equiv \frac{1}{u_1! \cdots u_d!} D^u q_k(\tau|x).$$

We construct an estimator $\hat{\gamma}_{\tau,k}(x)$ as follows. First, we define for each $\gamma \in \mathbf{R}^{|A_r|}$,

$$S_{n,x,\tau,k}(\gamma) \equiv \sum_{i=1}^n 1\{L_i = k\} \sum_{\ell=1}^{L_i} l_\tau \left[B_{\ell i} - \gamma^\top c \left(\frac{X_i - x}{h} \right) \right] K \left(\frac{x - X_i}{h} \right).$$

Then we construct

$$(2.1) \quad \hat{\gamma}_{\tau,k}(x) \equiv \operatorname{argmin}_{\gamma \in \mathbf{R}^{|A_r|}} S_{n,x,\tau,k}(\gamma),$$

where $l_\tau(u) \equiv u[\tau - 1\{u \leq 0\}]$ for any $u \in \mathbf{R}$, $K_h(t) = K(t/h)/h^d$, K is a d -variate kernel function, and h is a bandwidth that goes to zero as $n \rightarrow \infty$.

¹In fact, the estimator allows that we do not observe the whole vector B_i , but observe only B_{1i}, \dots, B_{ki} whenever $L_i = k$.

In order to reduce the redundancy of the statements, let us introduce the following definitions.

Definition 1. Let \mathcal{G} be a set of functions $g_v : \mathbf{R}^m \rightarrow \mathbf{R}^s$ indexed by a set \mathcal{V} , and let $S \subset \mathbf{R}^m$ be a given set and for $\varepsilon > 0$, let $S_v(\varepsilon)$ be an ε -enlargement of $S_v = \{x \in S : (x, v) \in S \times \mathcal{V}\}$, i.e., $S_v(\varepsilon) = \{x + a : x \in S \text{ and } a \in [-\varepsilon, \varepsilon]^m\}$. Then we define the following conditions for \mathcal{G} :

- (a) $B(S, \varepsilon)$: g_v is bounded on $S_v(\varepsilon)$ uniformly over $v \in \mathcal{V}$.
- (b) $BZ(S, \varepsilon)$: g_v is bounded away from zero on $S_v(\varepsilon)$ uniformly over $v \in \mathcal{V}$.
- (c) $BD(S, \varepsilon, r)$: \mathcal{G} satisfies $B(S, \varepsilon)$ and g_v is r times continuously differentiable on $S_v(\varepsilon)$ with derivatives bounded on $S_v(\varepsilon)$ uniformly over $v \in \mathcal{V}$.
- (d) $BZD(S, \varepsilon, r)$: \mathcal{G} satisfies $BZ(S, \varepsilon)$ and g_v is r times continuously differentiable on $S_v(\varepsilon)$ with derivatives bounded on $S_v(\varepsilon)$ uniformly over $v \in \mathcal{V}$.
- (e) LC : g_v is Lipschitz continuous with Lipschitz coefficient bounded uniformly over $v \in \mathcal{V}$.

Let \mathcal{P} denote the collection of the potential joint distributions of $(B^\top, X^\top, L)^\top$ and define $\mathcal{V} = \mathcal{T} \times \mathcal{P}$, and for each $k \in \mathbb{N}_L$,

$$\begin{aligned}
 (2.2) \quad \mathcal{G}_q(k) &= \{q_k(\tau|\cdot) : (\tau, P) \in \mathcal{V}\}, \\
 \mathcal{G}_f(k) &= \{f_{\tau,k}(\cdot|\cdot) : (\tau, P) \in \mathcal{V}\}, \\
 \mathcal{G}_L(k) &= \{P\{L_i = k | X_i = \cdot\} : P \in \mathcal{P}\}, \text{ and} \\
 \mathcal{G}_f &= \{f(\cdot) : P \in \mathcal{P}\},
 \end{aligned}$$

where $f_{\tau,k}(0|x)$ being the conditional density of $B_{li} - q_k(\tau|X_i)$ given $X_i = x$ and $L_i = k$. Also, define

$$(2.3) \quad \mathcal{G}_{f,2}(k) = \{f_{\cdot,k}(\cdot|\cdot) : P \in \mathcal{P}\} \text{ and } \mathcal{G}_\gamma(k) = \{\gamma_{\cdot,k}(\cdot) : P \in \mathcal{P}\}.$$

In other words, $\mathcal{G}_{f,2}(k)$ is the class of conditional densities $f_{\tau,k}(\cdot|x)$ indexed by τ , x , and probabilities P , and $\mathcal{G}_\gamma(k)$ is the class of functions $\gamma_{\tau,k}(\cdot)$ indexed by τ and probabilities P . We make the following assumptions.

Assumption QR 1.

- (i) \mathcal{G}_f satisfies $BD(\mathcal{S}, \varepsilon, 1)$.
- (ii) For each $k \in \mathbb{N}_L$, $\mathcal{G}_f(k)$ and $\mathcal{G}_L(k)$ satisfy both $BD(\mathcal{S}, \varepsilon, 1)$ and $BZD(\mathcal{S}, \varepsilon, 1)$.
- (iii) For each $k \in \mathbb{N}_L$, $\mathcal{G}_q(k)$ satisfies $BD(\mathcal{S}, \varepsilon, r+1)$ for some $r \geq 1$.
- (iv) For each $k \in \mathbb{N}_L$, $\mathcal{G}_{f,2}(k)$ and $\mathcal{G}_\gamma(k)$ satisfy LC .

Assumptions QR1(i) and (iii) are standard assumptions used in the local polynomial approach where one approximates $q_k(\cdot|x)$ by a linear combination of its derivatives through Taylor expansion, except only that the approximation here is required to behave well uniformly over $P \in \mathcal{P}$. Assumption QR1(ii) is made to prevent the degeneracy of the asymptotic

linear representation of $\hat{\gamma}_{\tau,k}(x) - \gamma_{\tau,k}(x)$ that is uniform over $x \in \mathcal{S}_\tau(\varepsilon)$, $\tau \in \mathcal{T}$ and over $P \in \mathcal{P}$. Assumption QR1(iv) requires that the conditional density function of $B_{li} - q_k(\tau|X_i)$ given $X_i = x$ and $L_i = k$ and $\gamma_{\tau,k}(\cdot)$ behave smoothly as we perturb τ locally. This requirement is used to control the size of the function spaces indexed by τ , so that when the stochastic convergence of random sequences holds, it is ensured to hold uniformly in τ .

Let $\|\cdot\|$ denote the Euclidean norm throughout the paper. Assumption QR2 lists conditions for the kernel function and the bandwidth.

Assumption QR 2. (i) K is compact-supported, bounded, and Lipschitz continuous on the interior of its support, $\int K(u)du = 1$, and $\int K(u)\|u\|^2du > 0$.

(ii) As $n \rightarrow \infty$, $n^{-1/2}h^{-d/2}\log n + \sqrt{nh^d}h^{r+1}/\sqrt{\log n} \rightarrow 0$, with r in Assumption QR1(iii).

Assumption QR2 gives conditions for the kernel and the bandwidth. The condition for the bandwidth is satisfied if we take $h = Cn^{-s}$ for some constant C with $s > 0$ satisfying that $1/(d + 2(r + 1)) < s < 1/d$.

For any sequence of real numbers $b_n > 0$, and any sequence of random vectors Z_n , we say that $Z_n/b_n \rightarrow_P 0$, \mathcal{P} -uniformly, or $Z_n = o_P(b_n)$, \mathcal{P} -uniformly, if for any $a > 0$,

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}} P\{\|Z_n\| > ab_n\} = 0.$$

Similarly, we say that $Z_n = O_P(b_n)$, \mathcal{P} -uniformly, if for any $a > 0$, there exists $M > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}} P\{\|Z_n\| > Mb_n\} < a.$$

Below, we establish a uniform Bahadur representation of $\sqrt{nh^d}H(\hat{\gamma}_{\tau,k}(x) - \gamma_{\tau,k}(x))$, where $H = \text{diag}((h^{|u|})_{u \in A_r})$ is the $|A_r| \times |A_r|$ diagonal matrix. First we introduce some notation. We define

$$\begin{aligned} \Delta_{x,\tau,lk,i} &\equiv B_{li} - \gamma_{\tau,k}^\top(x)c(X_i - x), \\ c_{h,x,i} &\equiv c((X_i - x)/h), \text{ and } K_{h,x,i} \equiv K((X_i - x)/h). \end{aligned}$$

Let

$$M_{n,\tau,k}(x) \equiv k \int P\{L_i = k|X_i = x + th\} f_{\tau,k}(0|x + th) f(x + th) K(t) c(t) c^\top(t) dt.$$

Theorem 1. Suppose that Assumptions QR1-QR2 hold. Then, for each $k \in \mathbb{N}_L$,

$$\begin{aligned} &\sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \left\| \sqrt{nh^d}H(\hat{\gamma}_{\tau,k}(x) - \gamma_{\tau,k}(x)) - M_{n,\tau,k}^{-1}(x)(\psi_{n,x,\tau,k} - \mathbf{E}\psi_{n,x,\tau,k}) \right\| \\ &= O_P\left(\frac{\log^{1/2} n}{n^{1/4}h^{d/4}}\right), \text{ } \mathcal{P}\text{-uniformly,} \end{aligned}$$

where, with $\tilde{l}_\tau(x) \equiv \tau - 1\{x \leq 0\}$,

$$\psi_{n,x,\tau,k} \equiv -\frac{1}{\sqrt{nh^d}} \sum_{i=1}^n 1\{L_i = k\} \sum_{l=1}^{L_i} \tilde{l}_\tau(\Delta_{x,\tau,lk,i}) c_{h,x,i} K_{h,x,i}.$$

The proof in this paper uses the convexity arguments of Pollard (1991) (see Kato (2009) for a related recent result) and, similarly as in Guerre and Sabbah (2012), employs the maximal inequality of Massart (2007, Theorem 6.8). The theoretical innovation of Theorem 1 is that we have obtained an approximation that is uniform in (x, τ) as well as in P . See Remark 1 below for a detailed comparison.

Remark 1. The main difference between this paper and Guerre and Sabbah (2012) is that their result pays attention to uniformity in h over some range, while our result pays attention to uniformity in P . Also it is interesting to note that the error rate here is a slight improvement over theirs. When $d = 1$, the rate here is $O_P[\sqrt{\log n}/(n^{1/4}h^{1/4})]$ while the rate in Theorem 2 of Guerre and Sabbah (2012) is $O_P[\log^{3/4} n/(n^{1/4}h^{1/4})]$. The difference is due to our use of an improved inequality which leads to a tighter bound in the maximal inequality in deriving the uniform error rate. For details, see the remark after the proof of Theorem 1 in the appendix.

The summands in the asymptotic linear representation form in Theorem 1 depend on the sample size and are not centered conditional on X_i . While this form can be useful in some contexts, the form is less illuminating. We provide an asymptotic linear representation that ensures this conditional centering given X_i .

Corollary 1. *Suppose that Assumptions QR1-QR2 hold. Then, for each $k \in \mathbb{N}_L$,*

$$\begin{aligned} & \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \left\| \sqrt{nh^d} H(\hat{\gamma}_{\tau,k}(x) - \gamma_{\tau,k}(x)) - M_{n,\tau,k}^{-1}(x) \tilde{\psi}_{n,x,\tau,k} \right\| \\ &= O_P \left(\frac{\log^{1/2} n}{n^{1/4}h^{d/4}} \right), \text{ } \mathcal{P}\text{-uniformly,} \end{aligned}$$

where

$$\tilde{\psi}_{n,x,\tau,k} \equiv -\frac{1}{\sqrt{nh^d}} \sum_{i=1}^n 1\{L_i = k\} \sum_{l=1}^{L_i} \tilde{l}_\tau(\varepsilon_{\tau,lk,i}) c_{h,x,i} K_{h,x,i}.$$

Note that the quantity $M_{n,\tau,k}(x)$ in the representation can be replaced by

$$M_{\tau,k}(x) \equiv k \int P\{L_i = k | X_i = x\} f_{\tau,k}(0|x) f(x) K(t) c(t) c^\top(t) dt,$$

if we modify the conditions on kernels and the smoothness conditions for the nonparametric function $P\{L_i = k|X_i = x\}f_{\tau,k}(0|x)f(x)$. As this modification can be done in a standard manner, we do not pursue details.

2.2. Bootstrap Uniform Bahadur Representation for Local Polynomial Quantile Regression Estimator. Let us consider the bootstrap version of the Bahadur representation in Theorem 1. Suppose that $\{(Y_i^{*\top}, X_i^{*\top})\}_{i=1}^n$ is a bootstrap sample drawn with replacement from the empirical distribution of $\{(Y_i^\top, X_i^\top)\}_{i=1}^n$. Throughout the paper, the bootstrap distribution P^* is viewed as the distribution of $(Y_i^*, X_i^*)_{i=1}^n$, conditional on $(Y_i, X_i)_{i=1}^n$, and let \mathbf{E}^* be expectation with respect to P^* .

We define the notion of uniformity in the convergence of distributions under P^* . For any sequence of real numbers $b_n > 0$, and any sequence of random vectors Z_n^* , we say that $Z_n^*/b_n \rightarrow_{P^*} 0$, \mathcal{P} -uniformly, or $Z_n^* = o_{P^*}(b_n)$, \mathcal{P} -uniformly, if for any $a > 0$,

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}} P \{P^* \{||Z_n^*|| > ab_n\} > a\} = 0.$$

Similarly, we say that $Z_n^* = O_{P^*}(b_n)$, \mathcal{P} -uniformly, if for any $a > 0$, there exists $M > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}} P \{P^* \{||Z_n^*|| > Mb_n\} > a\} < a.$$

For $z = (x, \tau) \in \mathcal{Z}$, define $\Delta_{x,\tau,lk,i}^* \equiv Y_{l,i}^* - \gamma_{\tau,k}^\top(x)c(X_i^* - x)$, and let $c_{h,x,i}^*$ and $K_{h,x,i}^*$ be $c_{h,x,i}$ and $K_{h,x,i}$ except that X_i is replaced by X_i^* . Then the following theorem gives the bootstrap version of Theorem 1.

Theorem 2. *Suppose that Assumptions QR1-QR2 hold. Then for each $k \in \mathbb{N}_J$,*

$$\begin{aligned} & \sup_{(x,\tau) \in \mathcal{X}_1 \times \mathcal{T}} \left\| \sqrt{nh^d} H(\hat{\gamma}_{\tau,k}^*(x) - \hat{\gamma}_{\tau,k}(x)) - M_{n,\tau,k}^{-1}(x) (\psi_{n,x,\tau,k}^* - \mathbf{E}^* \psi_{n,x,\tau,k}^*) \right\| \\ &= O_{P^*} \left(\frac{\log^{1/2} n}{n^{1/4} h^{d/4}} \right), \text{ } \mathcal{P}\text{-uniformly,} \end{aligned}$$

where $\psi_{n,x,\tau,k}^* \equiv -\frac{1}{\sqrt{nh^d}} \sum_{i=1}^n 1\{L_i = k\} \sum_{l=1}^k \tilde{l}_\tau(\Delta_{x,\tau,lk,i}^*) c_{h,x,i}^* K_{h,x,i}^*$.

Theorem 2 is obtained by using Le Cam's Poissonization Lemma (see Giné (1997, Proposition 2.5)) and following the proof of Theorem 1. The bootstrap version of Corollary 1 follows immediately from Theorem 2.

3. TESTING MONOTONICITY OF QUANTILE REGRESSION

This section considers testing monotonicity of quantile regression. We first state the testing problem formally, give the form of test statistic, verify regularity conditions, and present results of simple Monte Carlo experiments.

3.1. Testing Problem. Let $q(\tau|x)$ denote the τ -th quantile of Y conditional on $X = x$, where $\tau \in (0, 1)$ and X is a scalar random variable. In this subsection, we consider testing monotonicity of quantile regression. Define $g_\tau(x) \equiv \partial q(\tau|x)/\partial x$. The null hypothesis and the alternative hypothesis are as follows:

$$(3.1) \quad \begin{aligned} H_0 &: g_\tau(x) \leq 0 \text{ for all } (\tau, x) \in \mathcal{T} \times \mathcal{X} \text{ against} \\ H_1 &: g_\tau(x) > 0 \text{ for some } (\tau, x) \in \mathcal{T} \times \mathcal{X}, \end{aligned}$$

where \mathcal{X} is contained in the support of X and $\mathcal{T} \subset (0, 1)$. The null hypothesis states that the quantile functions are increasing on \mathcal{X} for all $\tau \in \mathcal{T}$, and the alternative hypothesis is the negation of the hypothesis. If \mathcal{T} is a singleton, then testing (3.1) amounts to testing monotonicity of quantile regression at a fixed quantile.

Suppose that $q(\tau|x)$ is continuously differentiable on \mathcal{X} for each $\tau \in \mathcal{T}$. Then one natural approach is to test the sign restriction of the derivative of $q(\tau|x)$. In other words, we develop a test of inequality restrictions using the local polynomial estimator of $\partial q(\tau|x)/\partial x$.

One may consider various other forms of monotonicity tests for quantile regression. For example, one might be interested in monotonicity of an interquartile regression function. More specifically, let $\tau_1 < \tau_2$ be chosen from $(0, 1)$ and write $\Delta g_{\tau_1, \tau_2}(x) \equiv g_{\tau_2}(x) - g_{\tau_1}(x)$. Then the null hypothesis and the alternative hypothesis of monotonicity of the interquartile regression function are as follows:

$$(3.2) \quad \begin{aligned} H_{0,\Delta} &: \Delta g_{\tau_1, \tau_2}(x) \leq 0 \text{ for all } x \in \mathcal{X} \text{ against} \\ H_{1,\Delta} &: \Delta g_{\tau_1, \tau_2}(x) > 0 \text{ for some } x \in \mathcal{X}. \end{aligned}$$

The null hypothesis states that the interquartile regression function $q(\tau_2|x) - q(\tau_1|x)$ is increasing on \mathcal{X} . This type of monotonicity can be used to investigate whether the income inequality (in terms of interquartile comparison) becomes severe as certain demographic variable X such as age increases.

3.2. Test Statistic. Suppose that we are given a random sample $\{(Y_i, X_i)\}_{i=1}^n$ of (Y, X) . First, we estimate $g_\tau(x)$ by local polynomial estimation to obtain, say, $\hat{g}_\tau(x) \equiv \mathbf{e}_2^\top \hat{\gamma}_\tau(x)$, where \mathbf{e}_2 is a column vector whose second entry is one and the rest zero, and

$$(3.3) \quad \hat{\gamma}_\tau(x) \equiv \operatorname{argmin}_{\gamma \in \mathbf{R}^{r+1}} \sum_{i=1}^n l_\tau(Y_i - \gamma^\top c(X_i - x)) K_h(X_i - x).$$

This paper applies the general approach of LSW, and proposes a new nonparametric test of monotonicity hypotheses involving quantile regression functions. First, testing the monotonicity of $q(\tau|\cdot)$ is tantamount to testing nonnegativity of g_τ on a domain of interest. Define $\Lambda_p(a) = (\max\{a, 0\})^p$ for any real numbers $a > 0$ and $p \geq 1$. We consider two test

statistics corresponding to the two testing problems discussed in Section 3.1: for $1 \leq p < \infty$,

$$(3.4) \quad \begin{aligned} T_{n,2} &\equiv \int_{\mathcal{X} \times \mathcal{T}} \Lambda_p(\hat{g}_\tau(x)) w_\tau(x) d(x, \tau), \text{ and} \\ T_n^\Delta &\equiv \int_{\mathcal{X}} \Lambda_p(\Delta \hat{g}_{\tau_1, \tau_2}(x)) w(x) dx, \quad \tau_1, \tau_2 \in (0, 1), \end{aligned}$$

where $\Delta \hat{g}_{\tau_1, \tau_2}(x) \equiv \hat{g}_{\tau_2}(x_1) - \hat{g}_{\tau_1}(x_1)$, $w_\tau(\cdot)$ and $w(\cdot)$ are nonnegative weight functions. The test statistic $T_{n,2}$ is used to test H_0 against H_1 and the test statistic $T_{n,\Delta}$ to test $H_{0,\Delta}$ against $H_{1,\Delta}$.

Let $\{(Y_i^*, X_i^*)\}$ be a bootstrap sample obtained from resampling from $\{(Y_i, X_i)\}_{i=1}^n$ with replacement. Then we define

$$(3.5) \quad \hat{\gamma}_\tau^*(x) \equiv \operatorname{argmin}_{\gamma \in \mathbf{R}^{r+1}} \sum_{i=1}^n l_\tau(Y_i^* - \gamma^\top c(X_i^* - x)) K_h(X_i^* - x)$$

and take $\hat{g}_\tau^*(x) \equiv \mathbf{e}_2^\top \hat{\gamma}_\tau^*(x)$, similarly as before. We construct the “recentered” bootstrap test statistics:

$$(3.6) \quad \begin{aligned} T_{n,2}^* &\equiv \int_{\mathcal{X} \times \mathcal{T}} \Lambda_p(\hat{g}_\tau^*(x) - \hat{g}_\tau(x)) w_\tau(x) d(x, \tau), \text{ and} \\ T_{n,\Delta}^* &\equiv \int_{\mathcal{X}} \Lambda_p(\Delta \hat{g}_{\tau_1, \tau_2}^*(x) - \Delta \hat{g}_{\tau_1, \tau_2}(x)) w(x) dx, \end{aligned}$$

where $\Delta \hat{g}_{\tau_1, \tau_2}^*(x) \equiv \hat{g}_{\tau_2}^*(x) - \hat{g}_{\tau_1}^*(x)$. We can now take the bootstrap critical values $c_{2,\alpha}^*$ and $c_{\Delta,\alpha}^*$ to be the $(1 - \alpha)$ quantiles from the bootstrap distributions of $T_{n,2}^*$ and $T_{n,\Delta}^*$. Then we define

$$c_{2,\alpha,\eta}^* = \max\{c_{2,\alpha}^*, h^{1/2}\eta + \hat{a}_2^*\} \text{ and } c_{\Delta,\alpha,\eta}^* = \max\{c_{\Delta,\alpha}^*, h^{1/2}\eta + \hat{a}_\Delta^*\},$$

where $\hat{a}_2^* \equiv \mathbf{E}^* T_{n,2}^*$ and $\hat{a}_\Delta^* \equiv \mathbf{E}^* T_{n,\Delta}^*$. The $(1 - \alpha)$ -level bootstrap tests for the two hypotheses are defined as

$$(3.7) \quad \begin{aligned} \text{Reject } H_0 &\quad \text{if and only if } T_{n,2} > c_{2,\alpha,\eta}^*. \\ \text{Reject } H_{0,\Delta} &\quad \text{if and only if } T_{n,\Delta} > c_{\Delta,\alpha,\eta}^*. \end{aligned}$$

3.3. Primitive Conditions. We present primitive conditions for the asymptotic validity of the proposed monotonicity tests. Let \mathcal{P} denote the collection of the potential joint distributions of $(Y, X)^\top$ and define $\mathcal{V} = \mathcal{T} \times \mathcal{P}$ as before. We also define $\tilde{\mathcal{V}} = \mathcal{T}^2 \times \mathcal{P}$ and define \mathcal{G}_f as in (2.2). Similarly as in (2.2), we introduce the following definitions:

$$\begin{aligned} \mathcal{G}_g &= \{g_\tau(\cdot) : (\tau, P) \in \mathcal{V}\}, \\ \mathcal{G}_{\Delta g} &= \{\Delta g_{\tau_1, \tau_2}(\cdot) : (\tau_1, \tau_2, P) \in \tilde{\mathcal{V}}\}, \text{ and} \\ \mathcal{G}_{Q,f} &= \{f_\tau(\cdot | \cdot) : (\tau, P) \in \mathcal{V}\}, \end{aligned}$$

where $f_\tau(0|x)$ being the conditional density of $Y_i - q(\tau|X_i)$ given $X_i = x$. Also, define

$$\mathcal{G}_{f,2} = \{f_\cdot(\cdot|\cdot) : P \in \mathcal{P}\} \text{ and } \mathcal{G}_\gamma = \{\gamma_\cdot(\cdot) : P \in \mathcal{P}\}.$$

We make the following assumptions.

Assumption MON 1. (i) \mathcal{G}_f satisfies $BD(\mathcal{S}, \varepsilon, 1)$.

(ii) $\mathcal{G}_{Q,f}$ satisfies both $BD(\mathcal{S}, \varepsilon, 1)$ and $BZD(\mathcal{S}, \varepsilon, 1)$.

(iii) \mathcal{G}_g satisfies $BD(\mathcal{S}, \varepsilon, r+1)$ for some $r > 3/2$.

(iv) $\mathcal{G}_{f,2}$ and \mathcal{G}_γ satisfy LC .

Assumption MON 2. (i) K is nonnegative and satisfies Assumption QR2(i).

(ii) $n^{-1/2}h^{-\{(3+\nu)/2\}+1} + \sqrt{nh^{r+2}}/\sqrt{\log n} \rightarrow 0$, as $n \rightarrow \infty$, for some small $\nu > 0$, with r in Assumption MON1(iii).

Assumption MON1 introduces a set of regularity conditions for various function spaces. Conditions (i)-(iii) require smoothness conditions. In particular, Condition (iii) is used to control the bias of the nonparametric quantile regression derivative estimator. Condition (iv) is analogous to Assumption QR1(iv) and used to control the size of the function space properly. Assumption MON2 introduces conditions for the kernel and bandwidth. The condition in Assumption MON2(ii) requires a bandwidth condition that is stronger than that in Assumption QR2(ii).

Assumptions IQM1-IQM2 below are used for testing $H_{0,\Delta}$.

Assumption IQM 1. (i) Assumptions MON1(i), (ii) and (iv) hold.

(ii) $\mathcal{G}_{\Delta g}$ satisfies $BD(\mathcal{S}, \varepsilon, r+1)$ for some $r > 3/2$.

Assumption IQM 2. The kernel function K and the bandwidth h satisfy Assumption MON2.

The following result establishes the uniform validity of the bootstrap test. Let $\mathcal{P}_0 \subset \mathcal{P}$ denote the set of potential distributions of the observed random vector under the null hypothesis.

Theorem 3. (i) Suppose that Assumptions MON1-MON2 hold. Then,

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}_0} P\{T_{n,2} > c_{2,\alpha,\eta}^*\} \leq \alpha.$$

(ii) Suppose that Assumptions IQM1-IQM2 hold. Then,

$$\limsup_{n \rightarrow \infty} \sup_{P \in \mathcal{P}_0} P\{T_{n,\Delta} > c_{\Delta,\alpha,\eta}^*\} \leq \alpha.$$

Using the general framework of LSW, it is possible to establish the consistency and local power of the test. Furthermore, it is also feasible to obtain a more powerful (but still asymptotically uniformly valid) test by estimating a contact set at the expense of requiring an additional tuning parameter (see LSW for details).

3.4. Monte Carlo Experiments. In this subsection, we present results of some Monte Carlo experiments that illustrate the finite-sample performance of one of the proposed tests. Specifically, we consider the following null and alternative hypotheses:

$$H_0 : \forall x \in \mathcal{X}, g(x) \geq 0 \text{ vs. } H_1 : \exists x \in \mathcal{X}, g(x) < 0,$$

where $g(x) \equiv \partial \text{Median}[Y|X=x] / \partial x$. In the experiments, X is generated independently from $\text{Unif}[0, 1]$ and U follows the distribution of $X^4 \times \mathbf{N}(0, 0.1)$.

To check the size of the test, we generated $Y = U$, which we call the null model. Note that the null model corresponds to the least favorable case. To see the power of the test, we considered the following alternative models: $Y = m_j(X) + U$ ($j = 1, 2, 3, 4, 5$), where

$$\begin{aligned} m_1(x) &= x(1-x), \\ m_2(x) &= -0.1x, \\ m_3(x) &= -0.1 \exp(-50(x-0.5)^2), \\ m_4(x) &= x + 0.6 \exp(-10x^2), \\ m_5(x) &= [10(x-0.5)^3 - 2 \exp(-10(x-0.5)^2)]1(x < 0.5) \\ &\quad + [0.1(x-0.5) - 2 \exp(-10(x-0.5)^2)]1(x \geq 0.5). \end{aligned}$$

In all experiments, $\mathcal{X} = [0.05, 0.95]$. Figures 1 and 2 show the true functions and corresponding simulated data.

The experiments use sample size of $n = 200$ and the nominal level of $\alpha = 0.10, 0.05$, and 0.01 . We performed 1,000 Monte Carlo replications for the null model and 200 replications for the alternative models. For each replication, we generated 200 bootstrap resamples. We used the local linear quantile regression estimator with the uniform kernel on $[-1/2, 1/2]$ for $K(\cdot)$. Furthermore, for the test statistic, we used $p = 2$ (one-sided L_2 norm) and uniform weight function $w(x) = 1$ and $h \in \{0.9, 1, 1.1\}$.

For the null model, the bootstrap approximation is quite good, especially with $h = 1$. The test shows good power for alternative models 1-3 and 5. For the alternative model 4, the power is sensitive with respect to the choice of the bandwidth. Overall, the finite sample behavior of the proposed test is satisfactory.

4. CONCLUSIONS

In this paper, we have established a uniform error rate of a Bahadur representation for the local polynomial quantile regression estimator. The error rate is uniform over a range of quantiles, a range of evaluation points in the regressors, and over a wide class of probabilities for observed random variables. We have illustrated the use of our Bahadur representation in the context of testing monotonicity. In addition, we have established the same error rate for the bootstrap local polynomial quantile regression estimator, which can be useful for bootstrap inference.

5. PROOFS

We also define for $a, b \in \mathbf{R}^{|A_r|}$,

$$\zeta_{n,x,\tau,k}(a, b) \equiv \sum_{i=1}^n 1\{L_i = k\} \sum_{l=1}^{L_i} \begin{Bmatrix} l_\tau \left(\Delta_{x,\tau,lk,i} - (a + b)^\top c_{h,x,i} / \sqrt{nh^d} \right) \\ -l_\tau \left(\Delta_{x,\tau,lk,i} - a^\top c_{h,x,i} / \sqrt{nh^d} \right) \end{Bmatrix} K_{h,x,i},$$

and

$$\zeta_{n,x,\tau,k}^\Delta(a, b) \equiv \zeta_{n,x,\tau,k}(a, b) - b^\top \psi_{n,x,\tau,k}.$$

Lemma QR1. *Suppose that Assumptions QR1-QR2 hold. Let $\{\delta_{1n}\}_{n=1}^\infty$ and $\{\delta_{2n}\}_{n=1}^\infty$ be positive sequences such that $\delta_{1n} = M\sqrt{\log n}$ for some $M > 0$ and $\delta_{2n} \leq \delta_{1n}$ from some large n on. Then for each $k \in \mathbb{N}_L$, the following holds uniformly over $P \in \mathcal{P}$:*

(i)

$$\begin{aligned} & \mathbf{E} \left[\sup_{a,b: \|a\| \leq \delta_{1n}, \|b\| \leq \delta_{2n}} \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} |\zeta_{n,x,\tau,k}^\Delta(a, b) - \mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a, b)]| \right] \\ &= O \left(\frac{\delta_{2n}\sqrt{\log n}}{n^{1/4}h^{d/4}} \right). \end{aligned}$$

(ii)

$$\mathbf{E} \left[\sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\psi_{n,x,\tau,k}\| \right] = O \left(\sqrt{\log n} + h^{r+1} \right) = O \left(\sqrt{\log n} \right).$$

(iii)

$$\begin{aligned} & \sup_{a,b: \|a\| \leq \delta_{1n}, \|b\| \leq \delta_{2n}} \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \left| \mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a, b)] - \frac{b^\top M_{n,\tau,k}(x)(b + 2a)}{2} \right| \\ &= O \left(\frac{\delta_{2n}\delta_{1n}^2}{n^{1/2}h^{d/2}} \right). \end{aligned}$$

Proof of Lemma QR1. (i) Define

$$(5.1) \quad \tilde{\delta}_{\tau,k}(x_1; x) \equiv q_k(\tau|x_1) - \gamma_{\tau,k}(x_1)^\top c(x_1 - x),$$

and

$$(5.2) \quad \delta_{n,\tau,k}(x_1; x) \equiv \tilde{\delta}_{\tau,k}(x_1; x) \mathbf{1}\{|x_1 - x| \leq h\},$$

where the dependence on P is through $q_k(\tau|x_1)$ and $\gamma_{\tau,k}(x_1)$. We also let

$$(5.3) \quad \delta_{n,\tau,k}(x_1) \equiv \sup_{x \in \mathcal{S}_\tau(\varepsilon)} \sup_{P \in \mathcal{P}} |\delta_{n,\tau,k}(x_1; x)|.$$

It is not hard to see that

$$(5.4) \quad \sup_{\tau \in \mathcal{T}, x_1 \in \mathcal{S}_\tau(\varepsilon)} |\delta_{n,\tau,k}(x_1)| = O(h^{r+1}),$$

because $q_k(\tau|x_1) - \gamma_{\tau,k}(x_1)^\top c(x_1 - x)$ is a residual from the Taylor expansion of $q_k(\tau|x_1)$ and \mathcal{X} is bounded, and the derivatives from the Taylor expansion are bounded uniformly over $P \in \mathcal{P}$.

Let $f_{\tau,k,x}^\Delta(t|x')$ be the conditional density of $\Delta_{x,\tau,lk,i}$ given $X_i = x'$. For all $x' \in \mathbf{R}^d$ such that $|x - x'| \leq h$,

$$(5.5) \quad \begin{aligned} f_{\tau,k,x}^\Delta(t|x') &= (\partial/\partial t)P\{B_{li} - q_k(\tau|X_i) \leq t - \delta_{n,\tau,k}(X_i; x)|X_i = x'\} \\ &= f_{\tau,k}(t - \delta_{n,\tau,k}(x'; x)|x'). \end{aligned}$$

Since $f_{\tau,k}(\cdot|x')$ is bounded uniformly over $x' \in \mathcal{S}_\tau(\varepsilon)$ and over $\tau \in \mathcal{T}$ (Assumption QR1(ii)), we conclude that for some $C > 0$ that does not depend on $P \in \mathcal{P}$,

$$(5.6) \quad \sup_{\tau \in \mathcal{T}} \sup_{x', x \in \mathcal{S}_\tau(\varepsilon)} \sup_{t \in \mathbf{R}} f_{\tau,k,x}^\Delta(t|x') < C.$$

We will use the results in (5.4) and (5.6) later.

Following the identity in Knight (1998, see the proof of Theorem 1), we write

$$l_\tau(x - y) - l_\tau(x) = -y \cdot \bar{l}_\tau(x) + \mu(x, y),$$

where $\mu(x, y) \equiv y \int_0^1 \{1\{x \leq ys\} - 1\{x \leq 0\}\} ds$ and

$$\bar{l}_\tau(x) \equiv \tau - 1\{x \leq 0\}.$$

Hence $\zeta_{n,x,\tau,k}^\Delta(a, b) - \mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a, b)]$ is equal to

$$\sum_{i=1}^n \{G_{n,x,\tau,k}(S_i; a, b) - \mathbf{E}[G_{n,x,\tau,k}(S_i; a, b)]\},$$

where $S_i \equiv (B_i^\top, X_i^\top, L_i)^\top$, $B_i = (B_{1,i}, \dots, B_{\bar{L},i})^\top$, and

$$(5.7) \quad G_{n,x,\tau}(S_i; a, b) \equiv \int_0^1 g_{n,x,\tau,k}(S_i; s, b, a) ds$$

and $g_{n,x,\tau,k}(S_i; s, b, a)$ is defined to be

$$1\{L_i = k\} \sum_{l=1}^k \begin{pmatrix} 1 \left\{ \Delta_{x,\tau,lk,i} - a^\top c_{h,x,i} / \sqrt{nh^d} \leq (sb)^\top c_{h,x,i} / \sqrt{nh^d} \right\} \\ -1 \left\{ \Delta_{x,\tau,lk,i} - a^\top c_{h,x,i} / \sqrt{nh^d} \leq 0 \right\} \end{pmatrix} \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}}.$$

Let $\mathcal{G}_n \equiv \{G_{n,x,\tau,k}(\cdot; a, b) : (a, b, x) \in [-\delta_{1n}, \delta_{1n}]^{r+1} \times [-\delta_{2n}, \delta_{2n}]^{r+1} \times \mathcal{S}_\tau(\varepsilon), \tau \in \mathcal{T}\}$,

$$\begin{aligned} \mathcal{G}_{1n} &\equiv \{\lambda_{\tau,1n}(\cdot; a, x) : (a, x) \in [-\delta_{1n}, \delta_{1n}]^{r+1} \times \mathcal{S}_\tau(\varepsilon), \tau \in \mathcal{T}\} \\ \mathcal{G}_{2n} &\equiv \{b^\top \lambda_{\tau,2n}(\cdot; x) : (b, x) \in [-\delta_{2n}, \delta_{2n}]^{r+1} \times \mathcal{S}_\tau(\varepsilon), \tau \in \mathcal{T}\} \text{ and} \\ \mathcal{G}_{3n} &\equiv \{\lambda_{\tau,3n}(\cdot; x) : x \in \mathcal{S}_\tau(\varepsilon), \tau \in \mathcal{T}\}, \end{aligned}$$

where $\lambda_{\tau,1n}(S_i; a, x) \equiv (\Delta_{x,\tau,lk,i} - a^\top \lambda_{\tau,2n}(S_i; x))_{l=1}^{\bar{L}}$,

$$\lambda_{\tau,2n}(S_i; x) \equiv c_{h,x,i} / \sqrt{nh^d} \text{ and } \lambda_{\tau,3n}(S_i; x) \equiv K_{h,x,i}.$$

First, we compute the entropy bound for \mathcal{G}_n . We focus on \mathcal{G}_{1n} first. By Assumption QR1(iv), there exists $C > 0$ that does not depend on $P \in \mathcal{P}$, such that for any $\tau \in \mathcal{T}$, any (a, x) and (a', x') in $[-\delta_{1n}, \delta_{1n}]^{r+1} \times \mathcal{S}_\tau(\varepsilon)$, and any $\tau, \tau' \in \mathcal{T}$,

$$|\lambda_{\tau,1n}(S_i; a, x) - \lambda_{\tau',1n}(S_i; a', x')| \leq Cn^s \{||a - a'|| + |\tau - \tau'| + ||x - x'||\},$$

for some $s > 0$. Observe that for any $\varepsilon' > 0$,

$$N(\varepsilon', [-\delta_{1n}, \delta_{1n}]^{r+1} \times \mathcal{S}_\tau(\varepsilon), ||\cdot||) \leq (\delta_{1n}/\varepsilon')^{r+2},$$

because $\mathcal{S}_\tau(\varepsilon)$ is bounded in the Euclidean space uniformly in $\tau \in \mathcal{T}$. Hence there are $C, s' > 0$ such that for all $\varepsilon' \in (0, 1]$,

$$\log N(\varepsilon', \mathcal{G}_{1n}, ||\cdot||_\infty) \leq -C \log \left((\varepsilon'/\delta_{1n}) n^{-s'} \right),$$

where $||\cdot||_\infty$ denotes the usual supremum norm. Applying similar arguments to \mathcal{G}_{2n} and \mathcal{G}_{3n} , we conclude that

$$(5.8) \quad \log N(\varepsilon', \mathcal{G}_{mn}, ||\cdot||_\infty) \leq C - C \log(\varepsilon'/n), \quad m = 1, 2, 3,$$

for some $C > 0$.

Define for $x \in \mathbf{R}$, $\delta > 0$,

$$\begin{aligned} 1_\delta^L(x) &\equiv (1 - \min\{x/\delta, 1\}) 1\{0 < x\} + 1\{x \leq 0\} \text{ and} \\ 1_\delta^U(x) &\equiv (1 - \min\{(x/\delta) + 1, 1\}) 1\{0 < x + \delta\} + 1\{x + \delta \leq 0\}. \end{aligned}$$

We also define for $x, y, z \in \mathbf{R}$,

$$\begin{aligned}\mu(x, y, z) &\equiv zy \int_0^1 \{1\{x \leq ys\} - 1\{x \leq 0\}\} ds, \\ \mu_\delta^U(x, y, z) &\equiv zy \int_0^1 \{1\{x \leq ys\} - 1_\delta^U(x)\} ds, \text{ and} \\ \mu_\delta^L(x, y, z) &\equiv zy \int_0^1 \{1\{x \leq ys\} - 1_\delta^L(x)\} ds.\end{aligned}$$

Then observe that

$$\begin{aligned}(5.9) \quad \mu_\delta^L(x, y, z) &\leq \mu(x, y, z) \leq \mu_\delta^U(x, y, z) \\ |\mu_\delta^U(x, y, z) - \mu(x, y, z)| &\leq |zy|1\{|x| < \delta\} \\ |\mu_\delta^L(x, y, z) - \mu(x, y, z)| &\leq |zy|1\{|x| < \delta\} \\ |\mu_\delta^U(x, y, z) - \mu_\delta^U(x', y', z')| &\leq C\{|y - y'| + |z - z'| + |x - x'|/\delta\}, \text{ and} \\ |\mu_\delta^L(x, y, z) - \mu_\delta^L(x', y', z')| &\leq C\{|y - y'| + |z - z'| + |x - x'|/\delta\},\end{aligned}$$

for any $y, y', x, x', z, z' \in \mathbf{R}$. Define

$$\begin{aligned}\mathcal{G}_{n,\delta}^U &\equiv \{\mu_\delta^U(g_1(S_i), g_2(S_i), g_3(S_i)) : g_m \in \mathcal{G}_{mn}, m = 1, 2, 3\}, \text{ and} \\ \mathcal{G}_{n,\delta}^L &\equiv \{\mu_\delta^L(g_1(S_i), g_2(S_i), g_3(S_i)) : g_m \in \mathcal{G}_{mn}, m = 1, 2, 3\}.\end{aligned}$$

From (5.9) and (5.8), we find that there exists $C > 0$ such that for each $\delta > 0$ and $\varepsilon > 0$,

$$\begin{aligned}(5.10) \quad \log N_{[]} (C\varepsilon, \mathcal{G}_{n,\delta}^U, L_p(P)) &\leq C - C \log(\varepsilon\delta/n) \text{ and} \\ \log N_{[]} (C\varepsilon, \mathcal{G}_{n,\delta}^L, L_p(P)) &\leq C - C \log(\varepsilon\delta/n).\end{aligned}$$

Fix $\varepsilon > 0$, set $\delta = \varepsilon$, and take brackets $[g_{1,L}^{(\varepsilon)}, g_{1,U}^{(\varepsilon)}], \dots, [g_{N,L}^{(\varepsilon)}, g_{N,U}^{(\varepsilon)}]$ and $[\tilde{g}_{1,L}^{(\varepsilon)}, \tilde{g}_{1,U}^{(\varepsilon)}], \dots, [\tilde{g}_{N,L}^{(\varepsilon)}, \tilde{g}_{N,U}^{(\varepsilon)}]$ such that

$$\begin{aligned}(5.11) \quad \mathbf{E} \left(|g_{s,U}^{(\varepsilon)}(S_i) - g_{s,L}^{(\varepsilon)}(S_i)|^2 \right) &\leq \varepsilon^2 \text{ and} \\ \mathbf{E} \left(|\tilde{g}_{s,U}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^2 \right) &\leq \varepsilon^2,\end{aligned}$$

and for any $g \in \mathcal{G}_n^U$ and $\tilde{g} \in \mathcal{G}_n^L$, there exists $s \in \{1, \dots, N\}$ such that $g_{s,L}^{(\varepsilon)} \leq g \leq g_{s,U}^{(\varepsilon)}$ and $\tilde{g}_{s,L}^{(\varepsilon)} \leq \tilde{g} \leq \tilde{g}_{s,U}^{(\varepsilon)}$. Without loss of generality, we assume that $g_{s,L}^{(\varepsilon)}, g_{s,U}^{(\varepsilon)} \in \mathcal{G}_n^U$ and $\tilde{g}_{s,L}^{(\varepsilon)}, \tilde{g}_{s,U}^{(\varepsilon)} \in \mathcal{G}_n^L$. By the first inequality in (5.9), we find that the brackets $[\tilde{g}_{s,L}^{(\varepsilon)}, g_{s,U}^{(\varepsilon)}]$, $k = 1, \dots, N$, cover \mathcal{G}_n . Hence by putting $\delta = \varepsilon$ in (5.10) and redefining constants, we conclude that for some $C > 0$

$$(5.12) \quad \log N_{[]} (C\varepsilon, \mathcal{G}_n, L_p(P)) \leq C - C \log(\varepsilon/n),$$

for all $\varepsilon > 0$.

Now, observe that

$$(5.13) \quad \sup_{b: \|b\| \leq \delta_{2n}, \tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \left| \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \right| \leq \frac{\bar{c} \|K\|_\infty \delta_{2n}}{\sqrt{nh^d}},$$

where $\bar{c} > 0$ is the diameter of the compact support of K .

For any $g \in \mathcal{G}_n/\bar{L}$ and any $m \geq 1$, we bound

$$(5.14) \quad |g(S_i)|^m \leq \left| \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \right|^m.$$

Also, for any $g \in \mathcal{G}_n/\bar{L}$, we use (5.7) and bound $\mathbf{E} [|g(S_i)|^2 | X_i, L_i = k]$ by

$$\left| \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \right|^2 P \left\{ - \left| \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \right| \leq \Delta_{x,\tau,lk,i} - \frac{a^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \leq \left| \frac{b^\top c_{h,x,i} K_{h,x,i}}{\sqrt{nh^d}} \right| | X_i, L_i = k \right\},$$

where $a \in [-\delta_{1n}, \delta_{1n}]^{r+1}$ and $b \in [-\delta_{2n}, \delta_{2n}]^{r+1}$ by the definition of \mathcal{G}_n . Using (5.13), we bound the last expression by

$$C_1 \frac{\delta_{2n}^3}{(nh^d)^{3/2}} \cdot \sup_{P \in \mathcal{P}} P \left\{ \max_{s=1,\dots,d} |X_{is} - x| \leq h/2 \right\} \leq C_2 \frac{\delta_{2n}^3 h^d}{(nh^d)^{3/2}},$$

for some constants $C_1, C_2 > 0$. Therefore, by (5.13), for some constants $C_1, C_2 > 0$, it is satisfied that for any $m \geq 2$,

$$\sup_{P \in \mathcal{P}} \mathbf{E} [|g(S_i)|^m] \leq C_1 \left(\frac{\delta_{2n}}{\sqrt{nh^d}} \right)^{m-2} \cdot \sup_{P \in \mathcal{P}} \mathbf{E} [|g(S_i)|^2] \leq C_2 b_n^{m-2} s_n^2,$$

where

$$(5.15) \quad b_n \equiv \frac{\delta_{2n}}{\sqrt{nh^d}} \text{ and } s_n \equiv \frac{\delta_{2n}^{3/2}}{n^{3/4} h^{d/4}}.$$

By (5.9), (5.11), and (5.13), and the definition of b_n and s_n in (5.15), there exist constants $C_1, C_2 > 0$ such that for all $m \geq 2$,

$$\begin{aligned} \mathbf{E} \left(|g_{s,U}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^m \right) &= \mathbf{E} \left(|g_{s,U}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^{m-2} |g_{s,U}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^2 \right) \\ &\leq C_1 \cdot b_n^{m-2} \cdot \mathbf{E} \left(|g_{s,U}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^2 \right) \\ &\leq 2C_1 \cdot b_n^{m-2} \cdot \mathbf{E} \left(|g_{s,U}^{(\varepsilon)}(S_i) - g_{s,L}^{(\varepsilon)}(S_i)|^2 \right) \\ &\quad + 2C_1 \cdot b_n^{m-2} \cdot \mathbf{E} \left(|g_{s,L}^{(\varepsilon)}(S_i) - \tilde{g}_{s,L}^{(\varepsilon)}(S_i)|^2 \right) \\ &\leq 2C_2 \cdot b_n^{m-2} \cdot \{\varepsilon^2 + b_n^2 \varepsilon\} \leq 2C_2 \cdot b_n^{m-2} \cdot \varepsilon. \end{aligned}$$

(The term $b_n^2 \varepsilon$ is obtained by chaining the second and third inequalities of (5.9) and using the fact that $\delta = \varepsilon$ and the uniform bound in (5.6). The last inequality follows because $b_n \rightarrow 0$ as $n \rightarrow \infty$.) We define $\bar{\varepsilon} = \varepsilon^{1/2}$ and bound the last term by $C_3 b_n^{m-2} \bar{\varepsilon}^2$, for some

$C_3 > 0$, because $b_n \leq 1$ from some large n on. The entropy bound in (5.12) as a function of $\bar{\varepsilon}$ remains the same except for a different constant $C > 0$ there.

Now by Theorem 6.8 of Massart (2007) and (5.12), there exist $C_1, C_2 > 0$ such that

$$\begin{aligned}
 (5.16) \quad & \sup_{P \in \mathcal{P}} \mathbf{E} \left[\sup_{a,b: \|a\| \leq \delta_{1n}, \|b\| \leq \delta_{2n}, \tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} |\zeta_{n,x,\tau,k}^\Delta(a, b) - \mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a, b)]| \right] \\
 & \leq C_1 \sqrt{n} \int_0^{s_n} \sqrt{n \wedge \left\{ -\log \left(\frac{\varepsilon}{n} \right) \right\}} d\varepsilon + C_1(b_n + s_n) \log n \\
 & \leq C_2 s_n \sqrt{n \log n} + C_2 b_n \log n = O \left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4} h^{d/4}} \right),
 \end{aligned}$$

where the last equality follows by the definitions of b_n and s_n in (5.15) and by Assumption QR2(ii).

(ii) Define $\lambda_{\tau,4n}(S_i; x) \equiv \Delta_{x,\tau,lk,i}$ and $\mathcal{L}_{k,1} \equiv \{\tilde{l}_\tau(\lambda_{\tau,4n}(\cdot; x)) : \tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)\}$, and $\mathcal{L}_{k,2} \equiv \{\lambda_{\tau,2n}(\cdot; x) \lambda_{\tau,3n}(\cdot; x) : \tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)\}$. We write

$$\psi_{n,x,\tau,k} = \{\psi_{n,x,\tau,k} - \mathbf{E}[\psi_{n,x,\tau,k}]\} + \mathbf{E}[\psi_{n,x,\tau,k}].$$

The leading term is an empirical process indexed by the functions in $\mathcal{L}_k \equiv \mathcal{L}_{k,1} \cdot \mathcal{L}_{k,2}$. Approximating the indicator function in \tilde{l}_τ by upper and lower Lipschitz functions and following similar arguments in the proof of (i), we find that

$$\sup_{P \in \mathcal{P}} \log N_{\mathbb{I}}(\varepsilon, \mathcal{L}_k, L_p(P)) \leq C - C \log \varepsilon + C \log n,$$

for some constant $C > 0$. Note that we can take a constant function C as an envelope of \mathcal{L}_k . Then we follow the proof of Lemma 2 to obtain that

$$\mathbf{E} \left[\sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\psi_{n,x,\tau,k} - \mathbf{E}[\psi_{n,x,\tau,k}]\| \right] = O(\sqrt{\log n}), \text{ uniformly in } P \in \mathcal{P}.$$

By using (5.4) and (5.5), we find that

$$(5.17) \quad \mathbf{E}[\psi_{n,x,\tau,k}] = O(h^{r+1}) = o(\sqrt{\log n}), \text{ uniformly in } P \in \mathcal{P},$$

because $\sqrt{nh^d}h^{r+1}/\sqrt{\log n} \rightarrow 0$ by Assumption QR2(ii).

(iii) Recall the definition of $g_{n,x,\tau,k}(S_i; s, b, a)$ in the proof of Lemma QR1(i). We write

$$\mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a, b)] = n \int_0^1 \mathbf{E}[g_{n,x,\tau,k}(S_i; s, b, a)] ds.$$

Using change of variables, we rewrite

$$\int_0^1 \mathbf{E}[g_{n,x,\tau,k}(S_i; s, b, a)] ds = kP\{L_i = k|X_i\} \cdot \phi_n(X_i; a, b),$$

where

$$\phi_n(X_i; a, b) = \int_{a^\top c_{h,x,i}/\sqrt{nh^d}}^{(b+a)^\top c_{h,x,i}/\sqrt{nh^d}} \begin{Bmatrix} F_{\tau,k}(u - \delta_{n,\tau,k}(X_i; x)|X_i) \\ -F_{\tau,k}(-\delta_{n,\tau,k}(X_i; x)|X_i) \end{Bmatrix} du \cdot K_{h,x,i}.$$

By expanding the difference, we have

$$\phi_n(X_i; a, b) = \int_{a^\top c_{h,x,i}/\sqrt{nh^d}}^{(b+a)^\top c_{h,x,i}/\sqrt{nh^d}} u du \cdot f_{\tau,k}(-\delta_{n,\tau,k}(X_i; x)|X_i) \cdot K_{h,x,i} + R_{n,x,i}(a, b),$$

where $R_{n,x,i}(a, b)$ denotes the remainder term in the expansion. As for the leading integral,

$$\int_{a^\top c_{h,x,i}/\sqrt{nh^d}}^{(b+a)^\top c_{h,x,i}/\sqrt{nh^d}} u du = \frac{1}{2nh^d} \{ b^\top c_{h,x,i} c_{h,x,i}^\top (b + 2a) \}.$$

Hence, for any sequences a_n, b_n , we can write $\mathbf{E}[\zeta_{n,x,\tau,k}^\Delta(a_n, b_n)]$ as

$$\begin{aligned} & \frac{1}{2} b_n^\top h^{-d} k \mathbf{E} [P\{L_i = k|X_i\} f_{\tau,k}(-\delta_{n,\tau,k}(X_i; x)|X_i) c_{h,x,i} c_{h,x,i}^\top \cdot K_{h,x,i}] (b_n + 2a_n) \\ & + nk \mathbf{E} [P\{L_i = k|X_i\} R_{n,x,i}(a_n, b_n)] \\ & = \frac{1}{2} b_n^\top M_{n,\tau,k}(x) (b_n + 2a_n) + nk \mathbf{E} [P\{L_i = k|X_i\} R_{n,x,i}(a_n, b_n)]. \end{aligned}$$

We can bound

$$\begin{aligned} & nk |\mathbf{E} [P\{L_i = k|X_i\} R_{n,x,i}(a_n, b_n)]| \\ & \leq C_1 nk \mathbf{E} \left[\int_{a_n^\top c_{h,x,i}/\sqrt{nh^d}}^{(b_n+a_n)^\top c_{h,x,i}/\sqrt{nh^d}} u^2 du \cdot K_{h,x,i} \right] \leq \frac{C_2 b_n a_n^2}{n^{1/2} h^{d/2}}, \end{aligned}$$

where $C_1 > 0$ and $C_2 > 0$ are constants that do not depend on n or $P \in \mathcal{P}$. \square

Proof of Theorem 1. (i) Let

$$\begin{aligned} (5.18) \quad \tilde{u}_{n,x,\tau} & \equiv -M_{n,\tau,k}^{-1}(x) \psi_{n,x,\tau,k}, \\ \tilde{\psi}_{n,x,\tau,k}(b) & \equiv b^\top \psi_{n,x,\tau,k} + b^\top M_{n,\tau,k}(x) b / 2, \text{ and} \\ \tilde{\psi}_{n,x,\tau,k}(a, b) & \equiv \tilde{\psi}_{n,x,\tau,k}(a + b) - \tilde{\psi}_{n,x,\tau,k}(a). \end{aligned}$$

For any $a \in \mathbf{R}^{|A_r|}$, we can write

$$\begin{aligned} (5.19) \quad \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}, a - \tilde{u}_{n,x,\tau}) & = \tilde{\psi}_{n,x,\tau,k}(a) - \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}) \\ & = (a - \tilde{u}_{n,x,\tau})^\top M_{n,\tau,k}(x) (a - \tilde{u}_{n,x,\tau}) / 2 \\ & \geq C_1 \|a - \tilde{u}_{n,x,\tau}\|^2, \end{aligned}$$

where $C_1 > 0$ is a constant that does not depend on $\tau \in \mathcal{T}$, $x \in \mathcal{S}_\tau(\varepsilon)$ or $P \in \mathcal{P}$. The last inequality uses Assumption QR1 and the fact that K is a nonnegative map that is not constant at zero and Lipschitz continuous.

Let

$$\hat{u}_{n,x,\tau} \equiv \sqrt{nh^d} H(\hat{\gamma}_{\tau,k}(x) - \gamma_{\tau,k}(x)),$$

where $x \in \mathcal{S}_\tau(\varepsilon)$ and $\tau \in \mathcal{T}$. Since $\zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, b)$ is convex in b , we have for any $0 < \delta \leq l$ and for any $b \in \mathbf{R}^{|\mathcal{A}_\tau|}$ such that $\|b\| = 1$,

$$\begin{aligned} (5.20) \quad (\delta/l) \zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, lb) &\geq \zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta b) \\ &\geq \tilde{\psi}_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta b) - \Delta_{n,k}(\delta), \end{aligned}$$

where

$$\Delta_{n,k}(\delta) \equiv \sup_{b \in \mathbf{R}^{|\mathcal{A}_\tau|}: \|b\|=1} |\zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta b) - \tilde{\psi}_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta b)|.$$

Therefore, if $\|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\| \geq \delta$, we replace b by $\hat{u}_{n,x,\tau}^\Delta = (\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau})/\|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\|$ and l by $\|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\|$ in (5.18), and use (5.20) to obtain that

$$\begin{aligned} (5.21) \quad 0 &\geq \zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\| \hat{u}_{n,x,\tau}^\Delta) \\ &\geq \zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta \hat{u}_{n,x,\tau}^\Delta) \\ &\geq \tilde{\psi}_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \delta \hat{u}_{n,x,\tau}^\Delta) - \Delta_{n,k}(\delta) \\ &\geq C_1 \delta^2 \|\hat{u}_{n,x,\tau}^\Delta\|^2 - \Delta_{n,k}(\delta) = C_1 \delta^2 - \Delta_{n,k}(\delta), \end{aligned}$$

for all $\delta \leq \|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\|$, where the first inequality follows because $\zeta_{n,x,\tau,k}(\hat{u}_{n,x,\tau}, \|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\| b)$ is minimized at $b = \hat{u}_{n,x,\tau}^\Delta$ by the definition of local polynomial estimation, the second and the third inequality follows by (5.20), and the fourth inequality follows from (5.19), and the last equality follows because $\|\hat{u}_{n,x,\tau}^\Delta\|^2 = 1$.

We take large $M > 0$ and let

$$(5.22) \quad \delta_{1n} = M \sqrt{\log n} \text{ and } \delta_{2n} = \frac{M \sqrt{\log n}}{n^{1/4} h^{d/4}}.$$

If $\delta_{2n} \leq \|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\|$, we have

$$C_1 \delta_{2n}^2 \leq \Delta_{n,k}(\delta_{2n}),$$

from (5.21). We let

$$1_n \equiv 1 \left\{ \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\tilde{u}_{n,x,\tau}\| \leq M \delta_{1n} \right\}.$$

Then we write

$$(5.23) \quad P \left\{ \inf_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\|^2 \geq \delta_{2n}^2 \right\} \leq P \{ \Delta_{n,k}(\delta_{2n}) 1_n \geq \delta_{2n}^2 \} + \mathbf{E}[1 - 1_n].$$

Now, we show that the first probability vanishes as $n \rightarrow \infty$. For each $b \in \mathbf{R}^{|A_r|}$, using the definition of $\tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}, b) = \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau} + b) - \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau})$, we write

$$\begin{aligned}\tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}, b) &= \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau} + b) - \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}) \\ &= b^\top \psi_{n,x,\tau,k} + (\tilde{u}_{n,x,\tau} + b)^\top M_{n,x,\tau}(\tilde{u}_{n,x,\tau} + b)/2 - \tilde{u}_{n,x,\tau}^\top M_{n,x,\tau} \tilde{u}_{n,x,\tau}/2 \\ &= b^\top \psi_{n,x,\tau,k} + b^\top M_{n,x,\tau} b/2 + b^\top M_{n,x,\tau} \tilde{u}_{n,x,\tau} \\ &= b^\top M_{n,x,\tau} b/2.\end{aligned}$$

Therefore,

$$\begin{aligned}\zeta_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}, b) - \tilde{\psi}_{n,x,\tau,k}(\tilde{u}_{n,x,\tau}, b) &= \zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b) - \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] \\ &\quad + \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] - b^\top M_{n,x,\tau} b/2 + b^\top \psi_{n,x,\tau,k} \\ &= \zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b) - \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] \\ &\quad + \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] - b^\top M_{n,x,\tau} (b + 2\tilde{u}_{n,x,\tau})/2.\end{aligned}$$

By Lemma QR1(i),

$$\begin{aligned}&\sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \sup_{b \in \mathbf{R}^{|A_r|}: \|b\| \leq \delta_{2n}} \left| \zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b) - \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] \right| \\ &= O_P \left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4} h^{d/4}} \right),\end{aligned}$$

by the definition in (5.22). And by Lemma QR1(iii),

$$\begin{aligned}&\sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \sup_{b \in \mathbf{R}^{|A_r|}: \|b\| \leq \delta_{2n}} \left| \mathbf{E} [\zeta_{n,x,\tau,k}^\Delta(\tilde{u}_{n,x,\tau}, b)] - b^\top M_{n,x,\tau} (b + 2\tilde{u}_{n,x,\tau})/2 \right| \\ &= O \left(\frac{\delta_{2n} \log n}{n^{1/2} h^{d/2}} \right),\end{aligned}$$

by the definition in (5.22) and Assumption QR2(ii). Thus we conclude that

$$(5.24) \quad |\Delta_{n,k}(\delta_{2n})| = O_P \left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4} h^{d/4}} \right),$$

where the last O_P term is uniform over $P \in \mathcal{P}$. We deduce from (5.24) that

$$\sup_{P \in \mathcal{P}} P \left\{ \Delta_{n,k}(\delta_{2n}) \mathbf{1} \left\{ \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\tilde{u}_{n,x,\tau}\| \leq \delta_{1n} \right\} \geq \delta_{2n}^{5/2} \right\} \rightarrow 0 \text{ as } n \rightarrow \infty$$

and as $M \uparrow \infty$. The proof is completed because

$$\sup_{P \in \mathcal{P}} P \left\{ \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \|\tilde{u}_{n,x,\tau}\| > \delta_{1n} \right\} \rightarrow 0,$$

as $n \rightarrow \infty$ and as $M \uparrow \infty$ by Lemma QR1(ii). Thus, we conclude from (5.23) that

$$\|\hat{u}_{n,x,\tau} - \tilde{u}_{n,x,\tau}\| = O_P\left(\frac{\sqrt{\log n}}{n^{1/4}h^{d/4}}\right), \text{ } \mathcal{P}\text{-uniformly.}$$

Now the desired result of Theorem 1 follows from the fact that

$$M_{n,x,\tau}^{-1} \mathbf{E} \psi_{n,x,\tau,k} = O(h^{r+1}) = o\left(\frac{\log^{1/2} n}{n^{1/4}h^{d/4}}\right),$$

which follows by Assumption QR1, (5.17), and Assumption QR2. \square

As mentioned in the main text, the convergence rate in the asymptotic linear representation is slightly faster than the rate in Theorem 2 of Guerre and Sabbah (2012). To see this difference closely, Guerre and Sabbah (2012) on page 118 wrote, for fixed numbers x and y ,

$$l_\tau(\varepsilon_i + x + y) - l_\tau(\varepsilon_i + x) - y \cdot \tilde{l}_\tau(\varepsilon_i + x) = \int_x^{x+y} (1\{\varepsilon_i \leq t\} - 1\{\varepsilon_i \leq 0\}) dt,$$

where ε_i is a certain random variable with density function, say, f which satisfies $\|f\|_\infty < \infty$. From this, Guerre and Sabbah (2012) proceeded as follows:

$$\begin{aligned} & \mathbf{E} \left[\left(l_\tau(\varepsilon_i + x + y) - l_\tau(\varepsilon_i + x) - y \cdot \tilde{l}_\tau(\varepsilon_i + x) \right)^2 \right] \\ & \leq 2|y| \int_x^{x+y} \mathbf{E} \left[(1\{\varepsilon_i \leq t\} - 1\{\varepsilon_i \leq 0\})^2 \right] dt \\ & \leq 2|y| \|f\|_\infty \int_x^{x+y} |t| dt \leq 2|y|^2 (|x| + |y|) \|f\|_\infty. \end{aligned}$$

On the other hand, this paper considers Knight (1998)'s inequality and proceeds as follows:

$$\begin{aligned} & \mathbf{E} \left[\left(l_\tau(\varepsilon_i + x + y) - l_\tau(\varepsilon_i + x) - y \cdot \tilde{l}_\tau(\varepsilon_i + x) \right)^2 \right] \\ & \leq |y|^2 \int_0^1 \mathbf{E} |1\{\varepsilon_i + x \leq yt\} - 1\{\varepsilon_i + x \leq 0\}| dt \\ & \leq |y|^2 P\{-|y| \leq \varepsilon_i + x \leq |y|\} \leq 2\|f\|_\infty |y|^3. \end{aligned}$$

Note that when $|y|$ is decreasing to zero faster than $|x|$, the latter bound is an improved one. The tighter L^2 bound gives a sharper bound when we apply the maximal inequality of Massart (2007) which yields a slightly faster error rate. (Compare Proposition A.1 of Guerre and Sabbah (2012) with Lemma QR1 where δ_{1n} and δ_{2n} in Lemma QR1 correspond to t_β and t_ϵ in Proposition A.1 respectively.)

Proof of Corollary 1. First, we write

$$M_{n,\tau,k}^{-1}(x)\psi_{n,x,\tau,k} = M_{n,\tau,k}^{-1}(x) \left(\tilde{\psi}_{n,x,\tau,k} + \frac{1}{\sqrt{nh^d}} \sum_{i=1}^n (\rho_{n,i}(x, \tau) - \mathbf{E}[\rho_{n,i}(x, \tau)]) \right),$$

where

$$\rho_{n,i}(x, \tau) = 1\{L_i = k\} \sum_{l=1}^{L_i} \left(\tilde{l}_\tau(\Delta_{x,\tau,lk,i}) - \tilde{l}_\tau(\varepsilon_{\tau,lk,i}) \right) c_{h,x,i} K_{h,x,i}.$$

It suffices for Corollary 1 to show that

$$M_{n,\tau,k}^{-1}(x) \frac{1}{\sqrt{nh^d}} \sum_{i=1}^n (\rho_{n,i}(x, \tau) - \mathbf{E}[\rho_{n,i}(x, \tau)]) = O_P \left(\frac{\log^{1/2} n}{n^{1/4} h^{d/4}} \right).$$

Using the definition in (5.1), writing $\tilde{\delta}_{x,\tau,k,i} = \tilde{\delta}_{\tau,k}(X_i; x)$, and using Knight's identity, we write

$$\begin{aligned} \rho_{n,i}(x, \tau) &= 1\{L_i = k\} \sum_{l=1}^{L_i} \left(\tilde{l}_\tau(\varepsilon_{\tau,lk,i} + \tilde{\delta}_{\tau,k}(X_i; x)) - \tilde{l}_\tau(\varepsilon_{\tau,lk,i}) \right) c_{h,x,i} K_{h,x,i} \\ &= 1\{L_i = k\} \tilde{\delta}_{x,\tau,k,i} \sum_{l=1}^{L_i} \left(\int_0^1 \begin{pmatrix} 1\{\varepsilon_{\tau,lk,i} \leq -\tilde{\delta}_{x,\tau,k,i} s\} \\ -1\{\varepsilon_{\tau,lk,i} \leq 0\} \end{pmatrix} ds \right) c_{h,x,i} K_{h,x,i}. \end{aligned}$$

Following the same arguments in the proof of Lemma QR1(i), we deduce that

$$M_{n,\tau,k}^{-1}(x) \frac{1}{\sqrt{nh^d}} \sum_{i=1}^n (\rho_{n,i}(x, \tau) - \mathbf{E}[\rho_{n,i}(x, \tau)]) = O_P \left(\frac{\log^{1/2} n}{n^{1/4} h^{d/4}} \right),$$

uniformly over $\tau \in \mathcal{T}$, $x \in \mathcal{S}_\tau(\varepsilon)$ and over $P \in \mathcal{P}$. \square

For $z = (x, \tau) \in \mathcal{Z}$ and $a, b \in \mathbf{R}$, we define

$$\zeta_{n,x,\tau,k}^*(a, b) \equiv \sum_{i=1}^n 1\{L_i = k\} \sum_{l=1}^k \begin{Bmatrix} l_\tau(\Delta_{x,\tau,lk,i}^* - (a+b)^\top c_{h,x,i}^*/\sqrt{nh^d}) \\ -l_\tau(\Delta_{x,\tau,lk,i}^* - b^\top c_{h,x,i}^*/\sqrt{nh^d}) \end{Bmatrix} K_{h,x,i}^*.$$

We also define

$$\zeta_{n,x,\tau,k}^{\Delta^*}(a, b) \equiv \zeta_{n,x,\tau,k}^*(a, b) - b^\top \psi_{n,x,\tau,k}^*.$$

The following lemma is the bootstrap analogue of Lemma QR1.

Lemma QR2. *Suppose that Assumptions QR1-QR2 hold. Let $\{\delta_{1n}\}_{n=1}^\infty$ and $\{\delta_{2n}\}_{n=1}^\infty$ be positive sequences such that $\delta_{1n} = M\sqrt{\log n}$ for some $M > 0$ and $\delta_{2n} \leq \delta_{1n}$ from some large n on. Then for each $k \in \mathbb{N}_L$, the following holds uniformly over $P \in \mathcal{P}$:*

(i)

$$\begin{aligned}
& \mathbf{E}^* \left[\sup_{a,b:||a|| \leq \delta_{1n}, ||b|| \leq \delta_{2n}} \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} |\zeta_{n,x,\tau,k}^{\Delta*}(a, b) - \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)]| \right] \\
&= O_P \left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4} h^{d/4}} \right).
\end{aligned}$$

(ii)

$$\begin{aligned}
& \sup_{a,b:||a|| \leq \delta_{1n}, ||b|| \leq \delta_{2n}} \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} \left| \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)] - \frac{b^\top M_{n,\tau,k}(x)(b + 2a)}{2} \right| \\
&= O_P \left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4} h^{d/4}} \right).
\end{aligned}$$

Note that the convergence rate in Lemma QR1(ii) is slower than that in Lemma QR1(iii).

Proof of Lemma QR2. (i) Similarly as in the proof of Lemma QR1(i), we rewrite $\zeta_{n,x,\tau,k}^{\Delta*}(a, b) - \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)]$ as

$$\sum_{i=1}^n \{G_{n,x,\tau,k}(S_i^*; a, b) - \mathbf{E}[G_{n,x,\tau,k}(S_i^*; a, b)]\},$$

where $S_i^* = (Y_i^{*\top}, X_i^{*\top})^\top$. Let $\pi = (x, \tau, s, a, b)$ and $\Pi_n = \mathcal{S}(\varepsilon) \times \mathcal{T} \times [0, 1] \times [-\delta_{1n}, \delta_{1n}]^{r+1} \times [-\delta_{2n}, \delta_{2n}]^{r+1}$, where $\mathcal{S}(\varepsilon) = \{(x, \tau) \in \mathcal{X} \times \mathcal{T} : x \in \mathcal{S}_\tau(\varepsilon)\}$. Using Proposition 2.5 of Giné (1997),

$$\begin{aligned}
& \mathbf{E} \left[\mathbf{E}^* \left[\sup_{a,b:||a|| \leq \delta_{1n}, ||b|| \leq \delta_{2n}} \sup_{\tau \in \mathcal{T}, x \in \mathcal{S}_\tau(\varepsilon)} |\zeta_{n,x,\tau,k}^{\Delta*}(a, b) - \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)]| \right] \right] \\
& \leq C \mathbf{E} \left[\mathbf{E}_{N_i} \left(\sup_{\pi \in \Pi_n} \left| \sum_{i=1}^n (N_i - 1) \left\{ g_{n,x,\tau,k}(S_i; s, b, a) - \frac{1}{n} \sum_{i=1}^n g_{n,x,\tau,k}(S_i; s, b, a) \right\} \right| \right) \right],
\end{aligned}$$

where $\{N_i\}_{i=1}^n$ are i.i.d. Poisson random variables with mean 1 independent of $\{(Y_i^\top, X_i^\top)^\top\}_{i=1}^\infty$, \mathbf{E}_{N_i} denotes expectation only with respect to the distribution of $\{N_i\}_{i=1}^n$, and $g_{n,x,\tau,k}(\cdot; s, b, a)$ is as defined in the proof of Lemma QR1(i). Here the constant $C > 0$ does not depend on $P \in \mathcal{P}$. We can bound the above by

$$\begin{aligned}
& C \mathbf{E} \left[\sup_{\pi \in \Pi_n} \left| \sum_{i=1}^n (N_i - 1) (g_{n,x,\tau,k}(S_i; s, b, a) - \mathbf{E}[g_{n,x,\tau,k}(S_i; s, b, a)]) \right| \right] \\
& + C \mathbf{E} \left(\left| \sum_{i=1}^n (N_i - 1) \right| \right) \times \mathbf{E} \left(\sup_{\pi \in \Pi_n} \left| \frac{1}{n} \sum_{i=1}^n g_{n,x,\tau,k}(S_i; s, b, a) - \mathbf{E}[g_{n,x,\tau,k}(S_i; s, b, a)] \right| \right).
\end{aligned}$$

The leading expectation is bounded by $O(\delta_{2n}\sqrt{\log n}/(n^{1/4}h^{d/4}))$ similarly as in the proof of Lemma QR1(i). And the product of the two expectations in the second term is bounded by

$$\begin{aligned} & O(\sqrt{n}) \times \frac{1}{n} \mathbf{E} \left(\sup_{\pi \in \Pi_n} \left| \sum_{i=1}^n \{g_{n,x,\tau,k}(S_i; s, b, a) - \mathbf{E}[g_{n,x,\tau,k}(S_i; s, b, a)]\} \right| \right) \\ &= O\left(\delta_{2n}^{3/2} \sqrt{\log n}/(n^{3/4}h^{d/4})\right), \end{aligned}$$

where the constant $C > 0$ does not depend on $P \in \mathcal{P}$, and the last equality follows similarly as in the proof of Lemma QR1(i).

(ii) Note that

$$(5.25) \quad \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)] = \mathbf{E}^*[\zeta_{n,x,\tau,k}^{\Delta*}(a, b)] - \mathbf{E}[\zeta_{n,x,\tau,k}^{\Delta}(a, b)] + \mathbf{E}[\zeta_{n,x,\tau,k}^{\Delta}(a, b)].$$

The difference between the first two terms on the right hand side is

$$O_P\left(\frac{\delta_{2n}^{3/2} \sqrt{\log n}}{n^{1/4}h^{d/4}}\right),$$

uniformly in $P \in \mathcal{P}$, as we have seen in (i). We apply Lemma QR1(iii) to the last expectation in (5.25) to obtain the desired result. \square

Proof of Theorem 2. The proof is completed by using Lemma QR2 precisely in the same way as the proof of Theorem 1 used Lemma QR1. While the convergence rate in Lemma QR2(ii) is slower than that in Lemma QR1(iii), we obtain the same convergence rate in the bootstrap version of (5.24). Details are omitted. \square

Lemma MIQ 1. (i) Suppose that the conditions of Theorem 3(i) hold. Then Assumptions A1-A3, A5-A6, and B1-B4 in LSW hold with the following definitions: $J = 1$, $r_n \equiv \sqrt{nh^3}$,

$$\begin{aligned} v_{n,\tau}(x) &\equiv \mathbf{e}_2^\top \gamma_\tau(x), \text{ and} \\ \beta_{n,x,\tau}(Y_i, z) &\equiv -\tilde{l}_\tau(Y_i - \gamma_\tau^\top(x) \cdot H \cdot c(z)) \mathbf{e}_2^\top M_{n,\tau}^{-1}(x) c(z) K(z). \end{aligned}$$

(ii) Suppose that the conditions of Theorem 3(ii) hold. Then Assumptions A1-A3, A5-A6, and B1-B4 in LSW hold with the following definitions: $J = 1$, $r_n \equiv \sqrt{nh^3}$,

$$\begin{aligned} v_{n,\tau}(x) &\equiv \mathbf{e}_2^\top \{\gamma_{\tau_1}(x) - \gamma_{\tau_2}(x)\}, \text{ and} \\ \beta_{n,x,\tau}(Y_i, z) &\equiv \alpha_{n,x,\tau_1}(Y_i, z) - \alpha_{n,x,\tau_2}(Y_i, z), \end{aligned}$$

where the set \mathcal{T} in LSW is replaced by $\mathcal{T} \times \mathcal{T}$ here, and

$$\alpha_{n,x,\tau}(Y_i, z) \equiv -\tilde{l}_\tau(Y_i - \gamma_\tau^\top(x) \cdot H \cdot c(z)) \mathbf{e}_2^\top M_{n,\tau}^{-1}(x) c(z) K(z).$$

Proof of Lemma MIQ. (i) First, Assumption A1 in LSW follows from Theorem 1, with the error rate in the asymptotic linear representation fulfills the rate $o_P(h^{1/2})$ by the condition: $r > 3/2$. Assumption A2 follows because $\beta_{n,x,\tau}(Y_i, z)$ has a multiplicative component of $K(z)$ having a compact support. As for Assumption A3, we can use Lemma 2 in LSW in combinations of Lipschitz continuity of $f(\cdot|\cdot)$ and $\gamma(\cdot)$ to verify the assumption. As we take $\hat{\sigma}_{\tau,j}(x) = \hat{\sigma}_{\tau,j}^*(x) = 1$, Assumptions A5 and B3 are trivially satisfied with the choice of $\sigma_{n,\tau,j}(x) = 1$. Assumption A6(i) is satisfied because $\beta_{n,x,\tau,j}$ is bounded. Assumptions Assumption B1 follows by Lemma QR2, and Assumption B2 by Lemma 2 in LSW. Assumption B4 follows by Assumption MON2(ii). (ii) The proof is similar and details are omitted. \square

Proof of Theorem 3. The results follow from Theorem 1 from LSW combined with Lemma MIQ1. Details are omitted. \square

REFERENCES

ANDREWS, D. W. K., X. CHENG, AND P. GUGGENBERGER (2011): “Generic Results for Establishing the Asymptotic Size of Confidence Sets and Tests,” Discussion Paper 1813, Cowles Foundation.

ANDREWS, D. W. K., AND X. SHI (2013): “Inference Based on Conditional Moment Inequalities,” *Econometrica*, 81(2), 609–666.

ANDREWS, D. W. K., AND G. SOARES (2010): “Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection,” *Econometrica*, 78(1), 119–157.

CHAUDHURI, P. (1991): “Nonparametric Estimates of Regression Quantiles and Their Local Bahadur Representation,” *Annals of Statistics*, 19(2), 760–777.

GINÉ, E. (1997): “Lectures on Some Aspects of the Bootstrap,” in *Lectures on Probability Theory and Statistics*, ed. by P. Bernard, vol. 1665 of *Lecture Notes in Mathematics*, pp. 37–151. Springer Berlin Heidelberg.

GUERRE, E., AND C. SABBAH (2012): “Uniform Bias Study and Bahadur Representation for Local Polynomial Estimators of the Conditional Quantile Function,” *Econometric Theory*, 28, 87–129.

KATO, K. (2009): “Asymptotics for Argmin Processes: Convexity Arguments,” *Journal of Multivariate Analysis*, 100(8), 1816–1829.

KNIGHT, K. (1998): “Limiting Distributions for L_1 Regression Estimators under General Conditions,” *Annals of Statistics*, 26(2), 755–770.

KONG, E., O. LINTON, AND Y. XIA (2010): “Uniform Bahadur Representation for Local Polynomial Estimates of m-Regression and Its Application to the Additive Model,” *Econometric Theory*, 26, 1529–1564.

——— (2013): “Global Bahadur Representation for Nonparametric Censored Regression Quantiles and Its Applications,” *Econometric Theory*, 29, 941–968.

LEE, S., K. SONG, AND Y.-J. WHANG (2015): “Testing a General Class of Functional Inequalities,” arXiv working paper, arXiv:1311.1595.

MASSART, P. (2007): *Concentration Inequalities and Model Selection*. Springer-Verlag, Berlin Heidelberg.

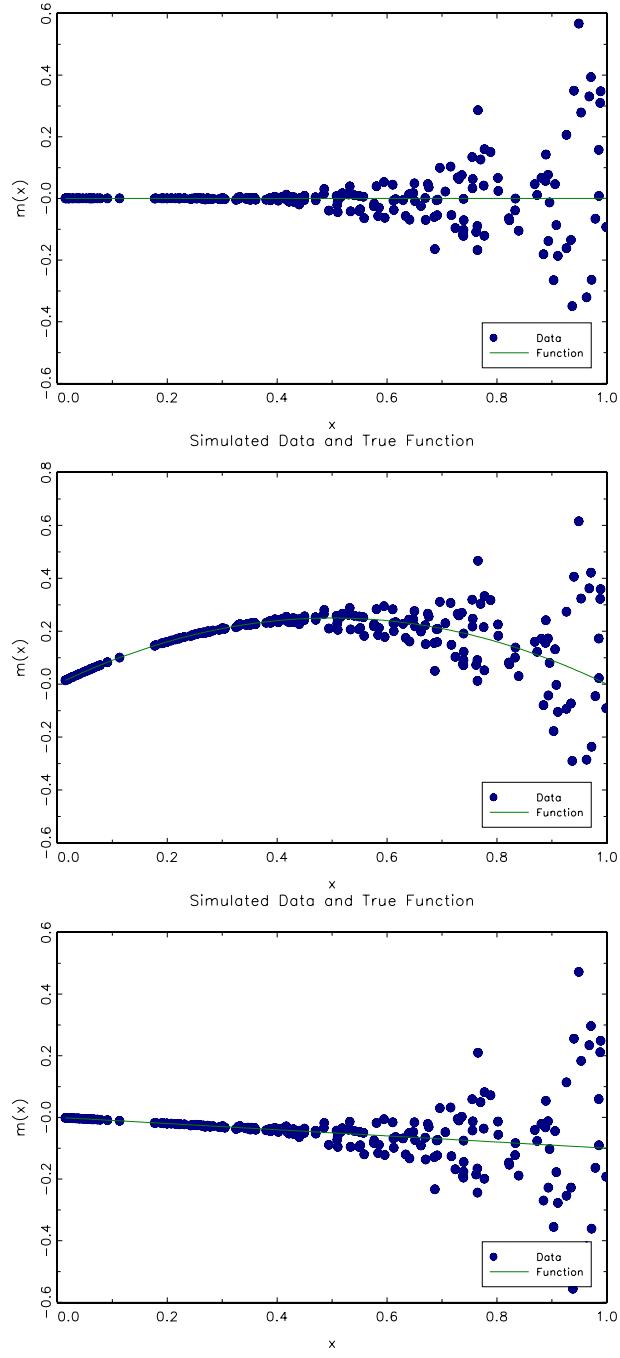
POLLARD, D. (1991): “Asymptotics for Least Absolute Deviation Regression Estimators,” *Econometric Theory*, 7, 186–199.

QU, Z., AND J. YOON (2015): “Nonparametric Estimation and Inference on Conditional Quantile Processes,” *Journal of Econometrics*, 185(1), 1–19.

TABLE 1. Results of Monte Carlo experiments

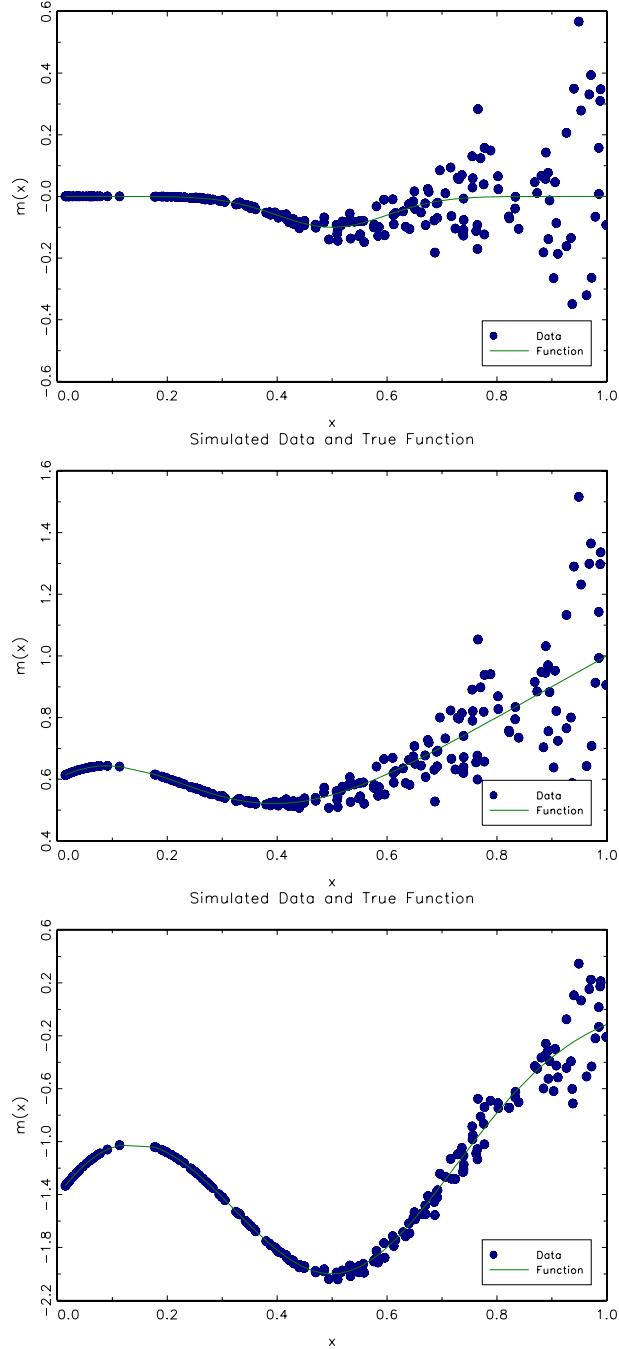
Bandwidth (h)	Null model			Alternative Model 1		
	Nominal level			Nominal level		
	0.10	0.05	0.01	0.10	0.05	0.01
0.9	0.111	0.057	0.020	0.995	0.975	0.780
1.0	0.100	0.048	0.007	0.980	0.920	0.660
1.1	0.077	0.036	0.005	0.905	0.755	0.375
Alternative Model 2						
Bandwidth (h)	Nominal level			Nominal level		
	0.10	0.05	0.01	0.10	0.05	0.01
	0.9	0.985	0.965	0.800	0.990	0.970
1.0	1.000	0.995	0.935	1.000	0.990	0.820
1.1	1.000	1.000	0.985	1.000	0.990	0.835
Alternative Model 4						
Bandwidth (h)	Nominal level			Nominal level		
	0.10	0.05	0.01	0.10	0.05	0.01
	0.9	1.000	0.960	0.540	0.995	0.980
1.0	0.885	0.645	0.175	0.995	0.995	0.985
1.1	0.310	0.120	0.010	0.995	0.990	0.935

FIGURE 1. True Function and Simulated Data
Simulated Data and True Function



Note: Each figure shows the true function and simulated data $\{(Y_i, X_i) : 1 = 1, \dots, n = 100\}$ being generated from $Y_i = m_j(X_i) + U_i$, where $X \sim \text{Unif}[0, 1]$ and $U_i \sim X^4 \times \mathbf{N}(0, 0.1^2)$, and $m_0(x) \equiv 0$, $m_1(x) = x(1-x)$, and $m_2(x) = -0.1x$, respectively.

FIGURE 2. True Function and Simulated Data
Simulated Data and True Function



Note: Each figure shows the true function and simulated data $\{(Y_i, X_i) : 1 = 1, \dots, n = 100\}$ being generated from $Y_i = m_j(X_i) + U_i$, where $X \sim \text{Unif}[0, 1]$ and $U_i \sim X^4 \times \mathcal{N}(0, 0.1^2)$, and $m_3(x) = -0.1 \exp(-50(x - 0.5)^2)$, $m_4(x) = x + 0.6 \exp(-10x^2)$, and $m_5(x) = [10((x - 0.5)^3) - 2 \exp(-10((x - 0.5)^2))]1(x < 0.5) + [0.1(x - 0.5) - 2 \exp(-10((x - 0.5)^2))]1(x \geq 0.5)$, respectively.